

RECEIVED

JUN 21 2001

TECH CENTER 1600/2900

1644

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/060,872A

DATE: 05/15/2001  
 TIME: 08:36:13

Input Set : A:\GC527-seqlist.txt  
 Output Set: N:\CRF3\05152001\I060872A.raw

3 <110> APPLICANT: Estell, David  
 4 Harding, Fiona  
 6 <120> TITLE OF INVENTION: MUTANT PROTEINS HAVING LOWER ALLERGENIC RESPONSE IN  
 7 HUMANS AND METHODS FOR CONSTRUCTING, IDENTIFYING AND  
 8 PRODUCING SUCH PROTEINS  
 10 <130> FILE REFERENCE: GC527  
 12 <140> CURRENT APPLICATION NUMBER: US 09/060,872A  
 13 <141> CURRENT FILING DATE: 1998-04-15  
 15 <160> NUMBER OF SEQ ID NOS: 211  
 17 <170> SOFTWARE: PatentIn Ver. 2.1  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 1495  
 21 <212> TYPE: DNA  
 22 <213> ORGANISM: Bacillus amyloliquefaciens  
 24 <220> FEATURE:  
 25 <221> NAME/KEY: mat\_peptide  
 26 <222> LOCATION: (417)..(1495)  
 28 <220> FEATURE:  
 29 <221> NAME/KEY: CDS  
 30 <222> LOCATION: (96)..(1244)  
 32 <220> FEATURE:  
 33 <221> NAME/KEY: misc\_feature  
 34 <222> LOCATION: (96)..(98)  
 35 <223> OTHER INFORMATION: The nnn at positions 96 through 98 represents gtg,  
 36 which is to code for methionine.  
 38 <220> FEATURE:  
 39 <221> NAME/KEY: misc\_feature  
 40 <222> LOCATION: (582)..(584)  
 41 <223> OTHER INFORMATION: The nnn at positions 582 through 584 represents  
 42 Xaa, which in a preferred embodiment (aat) is to  
 43 code for asparagine, but which may also code for  
 44 proline.  
 46 <220> FEATURE:  
 47 <221> NAME/KEY: misc\_feature  
 48 <222> LOCATION: (585)..(587)  
 49 <223> OTHER INFORMATION: The nnn at positions 585 through 587 represents  
 50 Xaa, which in a preferred embodiment (cct) is to  
 51 code for proline, but which may also code for  
 52 asparagine.  
 54 <220> FEATURE:  
 55 <221> NAME/KEY: misc\_feature  
 56 <222> LOCATION: (597)..(599)  
 57 <223> OTHER INFORMATION: The nnn at positions 597 to 599 represents Xaa,  
 58 which in a preferred embodiment (aac) is to code  
 59 for asparagine, but which may also code for  
 60 aspartic acid.  
 62 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 05/15/2001

PATENT APPLICATION: US/09/060,872A

TIME: 08:36:13

Input Set : A:\GC527-seqlist.txt

Output Set: N:\CRF3\05152001\I060872A.raw

63 <221> NAME/KEY: misc\_feature  
64 <222> LOCATION: (678)..(680)  
65 <223> OTHER INFORMATION: The nnn at positions 678 through 680 represents  
66 Xaa, which in a preferred embodiment (gca) is to  
67 code for alanine, but which may also code for  
68 serine.  
70 <220> FEATURE:  
71 <221> NAME/KEY: misc\_feature  
72 <222> LOCATION: (681)..(683)  
73 <223> OTHER INFORMATION: The nnn at positions 681 through 683 represents  
74 Xaa, which in a preferred embodiment (tca) is to  
75 code for serine, but which may also code for  
76 alanine.  
78 <220> FEATURE:  
79 <221> NAME/KEY: misc\_feature  
80 <222> LOCATION: (708)..(710)  
81 <223> OTHER INFORMATION: The nnn at positions 708 through 710 represents  
82 Xaa, which in a preferred embodiment (gct) is to  
83 code for alanine, but which may also code for  
84 aspartic acid.  
86 <220> FEATURE:  
87 <221> NAME/KEY: misc\_feature  
88 <222> LOCATION: (711)..(713)  
89 <223> OTHER INFORMATION: The nnn at positions 711 through 713 represents  
90 Xaa, which in a preferred embodiment (gac) is to  
91 code for aspartic acid, but which may also code  
92 for alanine.  
94 <220> FEATURE:  
95 <221> NAME/KEY: misc\_feature  
96 <222> LOCATION: (888)..(890)  
97 <223> OTHER INFORMATION: The nnn at positions 888 through 890 represents  
98 Xaa, which in a preferred embodiment (act) is to  
99 code for threonine, but which may also code for  
100 serine.  
102 <220> FEATURE:  
103 <221> NAME/KEY: misc\_feature  
104 <222> LOCATION: (891)..(893)  
105 <223> OTHER INFORMATION: The nnn at positions 891 through 893 represents  
106 Xaa, which in a preferred embodiment (tcc) is to  
107 code for serine, but which may also code for  
108 threonine.  
110 <220> FEATURE:  
111 <221> NAME/KEY: misc\_feature  
112 <222> LOCATION: (1167)..(1169)  
113 <223> OTHER INFORMATION: The nnn at positions 1167 through 1169 represents  
114 Xaa, which in a preferred embodiment (gaa) is to  
115 code for glutamic acid, but which may also code  
116 for glutamine.  
118 <400> SEQUENCE: 1

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/060,872A

DATE: 05/15/2001  
 TIME: 08:36:13

Input Set : A:\GC527-seqlist.txt  
 Output Set: N:\CRF3\05152001\I060872A.raw

W--> 119 ggtctactaa aatattattc catactatac aattaatāca cagaataatc tgtctattgg 60  
 121 ttattctgca aatgaaaaaa aggagaggat aaaga nnn aga ggc aaa aaa gta 113  
 122 Xaa Arg Gly Lys Lys Val  
 -105  
 123  
 125 tgg atc agt ttg ctg ttt gct tta gcg tta atc ttt acg atg gcg ttc 161  
 126 Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu Ile Phe Thr Met Ala Phe  
 -95 -90  
 127 -100  
 129 ggc agc aca tcc tct gcc cag gcg gca ggg aaa tca aac ggg gaa aag 209  
 130 Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys  
 -80 -75 -70  
 131 -85  
 133 aaa tat att gtc ggg ttt aaa cag aca atg agc acg atg agc gcc gct 257  
 134 Lys Tyr Ile Val Gly Phe Lys Gln Thr Met Ser Thr Met Ser Ala Ala  
 -65 -60 -55  
 135  
 137 aag aag aaa gat gtc att tct gaa aaa ggc ggg aaa gtg caa aag caa 305  
 138 Lys Lys Lys Asp Val Ile Ser Glu Lys Gly Gly Lys Val Gln Lys Gln  
 -50 -45 -40  
 139  
 141 ttc aaa tat gta gac gca gct tca gct aca tta aac gaa aaa gct gta 353  
 142 Phe Lys Tyr Val Asp Ala Ala Ser Ala Thr Leu Asn Glu Lys Ala Val  
 -35 -30 -25  
 143  
 145 aaa gaa ttg aaa aaa gac ccg agc gtc gct tac gtt gaa gaa gat cac 401  
 146 Lys Glu Leu Lys Lys Asp Pro Ser Val Ala Tyr Val Glu Glu Asp His  
 -20 -15 -10  
 147  
 149 gta gca cat gcg tac gcg cag tcc gtg cct tac ggc gta tca caa att 449  
 150 Val Ala His Ala Tyr Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile  
 -5 -1 1 5 10  
 151  
 153 aaa gcc cct gct ctg cac tct caa ggc tac act gga tca aat gtt aaa 497  
 154 Lys Ala Pro Ala Leu His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys  
 15 20 25  
 155  
 157 gta gcg gtt atc gac agc ggt atc gat tct tct cat cct gat tta aag 545  
 158 Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys  
 30 35 40  
 159  
 W--> 161 gta gca ggc gga gcc agc atg gtt cct tct gaa aca nnn nnn ttc caa 593  
 W--> 162 Val Ala Gly Gly Ala Ser Met Val Pro Ser Glu Thr Xaa Xaa Phe Gln  
 45 50 55  
 163  
 W--> 165 gac nnn aac tct cac gga act cac gtt gcc ggc aca gtt gcg gct ctt 641  
 W--> 166 Asp Xaa Asn Ser His Gly Thr His Val Ala Gly Thr Val Ala Ala Leu  
 60 65 70 75  
 167  
 W--> 169 aat aac tca atc ggt gta tta ggc gtt gcg cca agc nnn nnn ett tac 689  
 W--> 170 Asn Asn Ser Ile Gly Val Leu Gly Val Ala Pro Ser Xaa Xaa Leu Tyr  
 80 85 90  
 171  
 W--> 173 gct gta aaa gtt ctc ggt nnn nnn ggt tcc ggc caa tac agc tgg atc 737  
 W--> 174 Ala Val Lys Val Leu Gly Xaa Xaa Gly Ser Gly Gln Tyr Ser Trp Ile  
 95 100 105  
 175  
 177 att aac gga atc gag tgg gcg atc gca aac aat atg gac gtt att aac 785  
 178 Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn Asn Met Asp Val Ile Asn  
 110 115 120  
 179  
 181 atg agc ctc ggc gga cct tct ggt tct gct gct tta aaa gcg gca gtt 833  
 182 Met Ser Leu Gly Gly Pro Ser Gly Ser Ala Ala Leu Lys Ala Ala Val  
 125 130 135  
 183

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/060,872A

DATE: 05/15/2001  
TIME: 08:36:13

Input Set : A:\GC527-seqlist.txt  
Output Set: N:\CRF3\05152001\I060872A.raw

185 gat aaa gcc gtt gca tcc ggc gtc gta gtc gtt gcg gca gcc ggt aac 881  
 186 Asp Lys Ala Val Ala Ser Gly Val Val Val Val Ala Ala Ala Gly Asn 155  
 187 140 145 150 155  
 189 gaa ggc nnn nnn ggc agc tca agc aca gtg ggc tac cct ggt aaa tac 929  
 190 Glu Gly Xaa Xaa Gly Ser Ser Ser Thr Val Gly Tyr Pro Gly Lys Tyr 170  
 191 160 165 170  
 193 cct tct gtc att gca gta ggc gct gtt gac agc agc aac caa aga gca 977  
 194 Pro Ser Val Ile Ala Val Gly Ala Val Asp Ser Ser Asn Gln Arg Ala 185  
 195 175 180 185  
 197 tct ttc tca agc gta gga cct gag ctt gat gtc atg gca cct ggc gta 1025  
 198 Ser Phe Ser Ser Val Gly Pro Glu Leu Asp Val Met Ala Pro Gly Val 200  
 199 190 195 200  
 201 tct atc caa agc acg ctt cct gga aac aaa tac ggg gcg tac aac ggt 1073  
 202 Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly 215  
 203 205 210 215  
 205 acg tca atg gca tct ccg cac gtt gcc gga gcg gct gct ttg att ctt 1121  
 206 Thr Ser Met Ala Ser Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu 235  
 207 220 225 230 235  
 209 tct aag cac ccg aac tgg aca aac act caa gtc cgc agc agt tta nnn 1169  
 210 Ser Lys His Pro Asn Trp Thr Asn Thr Gln Val Arg Ser Ser Leu Xaa 250  
 211 240 245 250  
 213 aac acc act aca aaa ctt ggt gat tct ttc tac tat gga aaa ggg ctg 1217  
 214 Asn Thr Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu 265  
 215 255 260 265  
 217 atc aac gta cag gcg gca gct cag taa aacataaaaa accggccttg 1264  
 218 Ile Asn Val Gln Ala Ala Ala Gln 275  
 219 270 275  
 221 gccccgccgg tttttttatt ttttttctc cgcattgttca atccgctcca taatcgacgg 1324  
 223 atggctccct ctgaaaattt taacgagaaa cggcggttg acccggtcca gtcccgtaac 1384  
 225 ggccaagtcc tgaacgtct caatcgccgc ttcccggtt cgggtcagct caatgccgta 1444  
 227 acggtcggcg gcgttttct gataccggga gacggcattc gtaatcggtat c 1495  
 230 <210> SEQ ID NO: 2  
 231 <211> LENGTH: 382  
 232 <212> TYPE: PRT  
 233 <213> ORGANISM: Bacillus amyloliquefaciens  
 235 <220> FEATURE:  
 236 <221> NAME/KEY: VARIANT  
 237 <222> LOCATION: (1)...(382)  
 238 <223> OTHER INFORMATION: Xaa = Any Amino Acid  
 240 <400> SEQUENCE: 2  
 241 Xaa Arg Gly Lys Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu  
 242 1 5 10 15  
 243 Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly  
 244 20 25 30  
 245 Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met  
 246 35 40 45  
 247 Ser Thr Met Ser Ala Ala Lys Lys Lys Asp Val Ile Ser Glu Lys Gly  
 248 50 55 60  
 249 Gly Lys Val Gln Lys Gln Phe Lys Tyr Val Asp Ala Ala Ser Ala Thr

## RAW SEQUENCE LISTING

DATE: 05/15/2001

PATENT APPLICATION: US/09/060,872A

TIME: 08:36:13

Input Set : A:\GC527-seqlist.txt

Output Set: N:\CRF3\05152001\I060872A.raw

250 65 70 75 80  
 251 Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala  
 252 85 90 95  
 253 Tyr Val Glu Glu Asp His Val Ala His Ala Tyr Ala Gln Ser Val Pro  
 254 100 105 110  
 255 Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln Gly Tyr  
 256 115 120 125  
 257 Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser  
 258 130 135 140  
 259 Ser His Pro Asp Leu Lys Val Ala Gly Gly Ala Ser Met Val Pro Ser  
 260 145 150 155 160  
 W--> 261 Glu Thr Xaa Xaa Phe Gln Asp Xaa Asn Ser His Gly Thr His Val Ala  
 262 165 170 175  
 263 Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly Val Ala  
 264 180 185 190  
 W--> 265 Pro Ser Xaa Xaa Leu Tyr Ala Val Lys Val Leu Gly Xaa Xaa Gly Ser  
 266 195 200 205  
 267 Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn  
 268 210 215 220  
 269 Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly Pro Ser Gly Ser Ala  
 270 225 230 235 240  
 271 Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala Ser Gly Val Val Val  
 272 245 250 255  
 W--> 273 Val Ala Ala Ala Gly Asn Glu Gly Xaa Xaa Gly Ser Ser Ser Thr Val  
 274 260 265 270  
 275 Gly Tyr Pro Gly Lys Tyr Pro Ser Val Ile Ala Val Gly Ala Val Asp  
 276 275 280 285  
 277 Ser Ser Asn Gln Arg Ala Ser Phe Ser Ser Val Gly Pro Glu Leu Asp  
 278 290 295 300  
 279 Val Met Ala Pro Gly Val Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys  
 280 305 310 315 320  
 281 Tyr Gly Ala Tyr Asn Gly Thr Ser Met Ala Ser Pro His Val Ala Gly  
 282 325 330 335  
 283 Ala Ala Ala Leu Ile Leu Ser Lys His Pro Asn Trp Thr Asn Thr Gln  
 284 340 345 350  
 W--> 285 Val Arg Ser Ser Leu Xaa Asn Thr Thr Thr Lys Leu Gly Asp Ser Phe  
 286 355 360 365  
 287 Tyr Tyr Gly Lys Gly Leu Ile Asn Val Gln Ala Ala Gln  
 288 370 375 380  
 292 <210> SEQ ID NO: 3  
 293 <211> LENGTH: 275  
 294 <212> TYPE: PRT  
 295 <213> ORGANISM: Bacillus amyloliquefaciens  
 297 <400> SEQUENCE: 3  
 298 Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu  
 299 1 5 10 15  
 301 His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp  
 302 20 25 30  
 304 Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys Val Ala Gly Gly Ala

## VERIFICATION SUMMARY

DATE: 05/15/2001

\* PATENT APPLICATION: US/09/060,872A

TIME: 08:36:14

Input Set : A:\GC527-seqlist.txt

Output Set: N:\CRF3\05152001\I060872A.raw

L:121 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:122 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:166 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:170 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:173 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:174 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:190 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:209 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:261 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:265 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:285 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2